

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of

Agrawal et al.

Serial No.: 10/624,069

Group Art Unit: 2161

Filed: July 21, 2003

Examiner: Padmanabhan, Kavita

For: MINING ASSOCIATION RULES OVER PRIVACY PRESERVING DATA

Commissioner of Patents
P.O. BOX 1450
Alexandria, VA 22313-1450

DECLARATION UNDER 37 C.F.R. §1.132

We, Alexandre Evfimievski, Ramakrishnan Srikant, and Rakesh Agrawal, are the Applicants and joint inventors of the above-referenced invention disclosed in U.S. Patent Application Serial No. 10/624,069 and hereby declare the following:

[0001] Alexandre Evfimievski is the author of the following paper: "Randomization in Privacy Preserving Data Mining," December 2002, ACM SIGKDD Explorations Newsletter, Volume 4, Issue 2, pages 43-48, referred to herein as "Randomization in Privacy Preserving Data Mining" (December 2002) and attached (see Attachment A). This paper is the basis for a rejection of claims 1-24 of the above-referenced application under 35 U.S.C. 102(a).

[0002] "Randomization in Privacy Preserving Data Mining" (December 2002) is an overview of research in the field of privacy preserving data mining and is derived from several papers. Specifically, the "paper presents some methods and results in randomization for numerical and categorical data and discusses the issue of measuring privacy" (see Abstract).

[0003] A. Evfimievski derived portions of the paper "Randomization in Privacy Preserving Data Mining" (December 2002) directly from the following paper, which was published on July 23, 2002, in conjunction with The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining held July 23-26, 2002.

A. Evfimievski, R. Srikant, R. Agrawal and J. Gehrke, "Privacy Preserving Mining of Association Rules," Proc. Of 8th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD), July 2002, referred to herein as "Privacy Preserving Mining of Association Rules" (July 2002). (see Attachment B)

[0004] Each of the Applicants of the present invention are co-authors on the paper "Privacy Preserving Mining of Association Rules" (July 2002) along with J. Gehrke. "Privacy Preserving Mining of Association Rules" (July 2002) describes the invention defined by claims 1-24 and, specifically, was the paper on which the present application was based. The present application was filed within 1 year of the date of publication of "Privacy Preserving Mining of Association Rules" (July 2002).

[0005] J. Gehrke was a professor and advisor of A. Evfimievski, during the time period in which the idea for the invention was conceived. Although J. Gehrke is listed as a co-author of "Privacy Preserving Mining of Association Rules" (July 2002), he was not an inventor. It should be noted that although J. Gehrke's was not an inventor, his contributions to the paper were significant in that he acted as an advisor providing assistance in drafting the manuscript, in checking the manuscript for errors, in providing related literature for reference, etc. Additionally, J. Gehrke has read U.S. Patent Application Serial No. 10/624,069 and has declared that he is not an inventor of the invention defined by claims 1-24 (see Attachment C). We also acknowledge that J. Gehrke was not an inventor of the invention defined by claims 1-24. Therefore, the portions of "Privacy Preserving Mining of Association Rules" (July 2002) that describe the features of claims 1-24 describe the Applicants' own work and no one else's and further predate the cited prior art.

[0006] Regarding the rejected claims 1-24, the Office Action provides that the paper “Randomization in Privacy Preserving Data Mining” (December 2002) anticipates all of the various claimed features of the invention. However, many of the cited portions of “Randomization in Privacy Preserving Data Mining” (December 2002) are not the original work of A. Evfimievski, but rather were derived from the paper “Privacy Preserving Mining of Association Rules” (July 2002), which, as discussed above, describes the Applicants’ own work. Therefore, it necessarily follows that the cited portions of “Randomization in Privacy Preserving Data Mining” (December 2002), which the Office Action indicates as disclosing the features of claims 1-24, were also derived from the Applicants’ own work. Consequently, the cited portions of “Randomization in Privacy Preserving Data Mining” (December 2002) are not available as prior art against the present application.

[0007] More specifically, I derived the following portions of “Randomization in Privacy Preserving Data Mining” (December 2002) cited in the Office Action from the following portions of “Privacy Preserving Mining of Association Rules” (July 2002).

a. P. 43, left column, paragraph 3, lines 10 – right column, paragraph 1, line 2; p. 44, right column, paragraph 6, lines 2-4; and p. 45, left column, paragraph 4, lines 7-8 were each cited as disclosing both the feature of “randomly dropping true items from each transaction in said original dataset” and the feature of “randomly inserting false items into each transaction in said original data set.”

(1) Regarding p. 43, left column, paragraph 3, lines 10 – right column, paragraph 1, line 2, no specific reference is credited for the cited sentence. However, the idea of “before sending its piece of data, each client perturbs it so that some true information is taken away and some false information is introduced” is derived directly from the statement “... in addition to replacing some of the items, we shall insert so many “false” items into a transaction that one is as likely to see a “false itemset as a

“true” one”, which is found in the second sentence of section 4 of “Privacy Preserving Mining of Association Rules” (July 2002).

(2) Regarding p. 44, right column, paragraph 6, lines 2-4, no specific reference is credited for the cited sentence. The first sentence in paragraph 6 that immediately precedes the cited lines refers specifically to “Privacy Preserving Mining of Association Rules” (July 2002) along with another reference, the publication date of which is later (i.e., August 2002). Additionally, this sentence, “Suppose that each client C_i has a transaction t_i , which is a subset of a given finite set of items I , $|I| = n$ ” does not amount to “randomly dropping true items from each transaction in said original dataset” or “randomly inserting false items into each transaction in said original data set”, but rather defines the transaction prior to randomly dropping true items therefrom.

(3) Regarding p. 45, left column, paragraph 4, lines 7-8, again no specific reference is credited for the cited sentence; however, the sentence immediately following explains the statement and does credit “Privacy Preserving Mining of Association Rules” (July 2002). Additionally, the idea of a “natural way to randomize a set of items is by deleting some items and inserting some new items is derived directly from the statement “... in addition to replacing some of the items, we shall insert so many “false” items into a transaction that one is as likely to see a “false itemset as a “true” one” (see second sentence of section 4 of “Privacy Preserving Mining of Association Rules” (July 2002)).

b. P. 45, left column, paragraph 8, lines 1-3 is cited in the Office Action as disclosing the features of “creating a randomized dataset by collecting said randomized transactions” and of “collecting said randomized dataset in a database.” Again, no specific reference is credited for the cited sentence, which states “In the set T' of randomized transactions available to the server, itemsets have supports very different from their supports in the nonrandomized dataset T .” However, this sentence is contained within a section of the paper that comprises a summary of the article “Privacy Preserving Mining of Association Rules” (July 2002) and, specifically, a summary of

section 4 of the article. More specifically, the idea of “creating a randomized dataset by collecting said randomized transactions” is derived directly from Definition 5 of Section 4.1 of “Privacy Preserving Mining of Association Rules” (July 2002). That is, randomization R is “a per-transaction randomization if, for $T = (t_1, t_2, \dots, t_N)$, we can represent $R(T)$ as $R(t_1, t_2, \dots, t_N) = (R(1, t_1), R(2, t_2), \dots, R(N, t_N))$, where $R(i, t)$ are independent random variables whose distributions depend only on t (and not on i). We shall write $t'_i = R(i, t_i) = R(t_i)$.” It should be noted that the cited sentence on p. 45, left column, paragraph 8, lines 1-3 does not refer at all to collecting a randomized dataset in a database.

c. P. 45, left column, paragraph 9- right column, paragraph 6, line 9 is cited in the Office Action as teaching the feature of “mining said database to recover an association rule after said dropping and inserting processes by estimating nonrandomized support of said association rule in said original dataset based on randomized support for said association rule in said randomized dataset, wherein due to said creating of said randomized transactions, privacy breaches of said individual transactions are controlled during said mining.” This portion of “Randomization in Privacy Preserving Data Mining” (December 2002) specifically references “Privacy Preserving Mining of Association Rules” (July 2002) and is derived from Sections 4.3 and 4.4.

d. P. 45, left column, paragraph 4, line 13- paragraph 8, line 11 is cited in the Office Action as teaching the feature of “wherein said process of creating randomized transactions comprises per transaction randomizing, such that randomizing operators are applied to each transaction independently.” This portion of “Randomization in Privacy Preserving Data Mining” (December 2002) specifically references “Privacy Preserving Mining of Association Rules” (July 2002) and is derived directly from Sections 4.1.

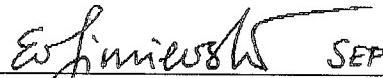
e. P. 45, left column, paragraph 4, line 7- paragraph 8, line 11 is cited in the Office Action as teaching the feature of “wherein said process of creating randomized transactions is item-invariant such that a reordering of said transactions does not affect outcome probabilities.” Again, this portion of “Randomization in Privacy Preserving

Data Mining" (December 2002) specifically references "Privacy Preserving Mining of Association Rules" (July 2002) and is derived directly from Remark 1 of Section 4.1.

f. P. 46, left column, paragraph 3, lines 1-13 is cited as teaching the features of "wherein said dropping of said true items and said inserting of said false items are carried out to an extent such that the chance of finding a false itemset in a randomized transaction relative to the chance of finding a true itemset in said randomized transaction is above a predetermined threshold" and "wherein said predetermined threshold provides that the chance of finding a false itemset in said randomized transaction is approximately equal to the chance of finding a true itemset in said randomized transaction." Again, this portion of "Randomization in Privacy Preserving Data Mining" (December 2002) specifically references "Privacy Preserving Mining of Association Rules" (July 2002) and is derived from the second sentence in Section 4 and Section 4.1.

g. P. 45, left column, paragraph 8, line 1- paragraph 9, line 33 is cited as teaching the feature of "wherein said process of creating randomized transactions is performed independently on said transactions prior to the transactions being collected in said database." No specific reference is credited for the cited sentence. However, this sentence is contained within a section of "Privacy Preserving Mining of Association Rules" (July 2002) that does refer to and summarize "randomization" as described in "Randomization in Privacy Preserving Data Mining" (December 2002).

[0008] The above declarations are made according to the best of our recollection upon review of the appropriate documents and notes, and we hereby acknowledge that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of the application or any patent issuing thereon. All statements that are made herein of our own knowledge are true and all statements that are made herein based on information and belief are believed to be true.


Alexandre Evgimievski SEPTEMBER 7, 2006
(Date)

Ramakrishnan Srikant (Date)

Rakesh Agrawal (Date)

[0008] The above declarations are made according to the best of our recollection upon review of the appropriate documents and notes, and we hereby acknowledge that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of the application or any patent issuing thereon. All statements that are made herein of our own knowledge are true and all statements that are made herein based on information and belief are believed to be true.

Alexandre Evfimievski (Date)

R. S. S. Sep 1, 2006
Ramakrishnan Srikant (Date)

Rakesh Agrawal (Date)